

Wavelets, Filter Banks, and the JPEG-2000 Image Coding Standard

Chris Brislawn, Los Alamos National Lab

Collaborators:

Sue Mniszewski, Allon Percus, Michelle Quirk, Reid Rivenburgh, Brendt Wohlberg (LANL)

Oliver Treiber (Univ. MD, IBM-Munich)

Bernie Brower (Eastman Kodak)

<http://www.c3.lanl.gov/~brislawn>, brislawn@lanl.gov

December 6, 2002

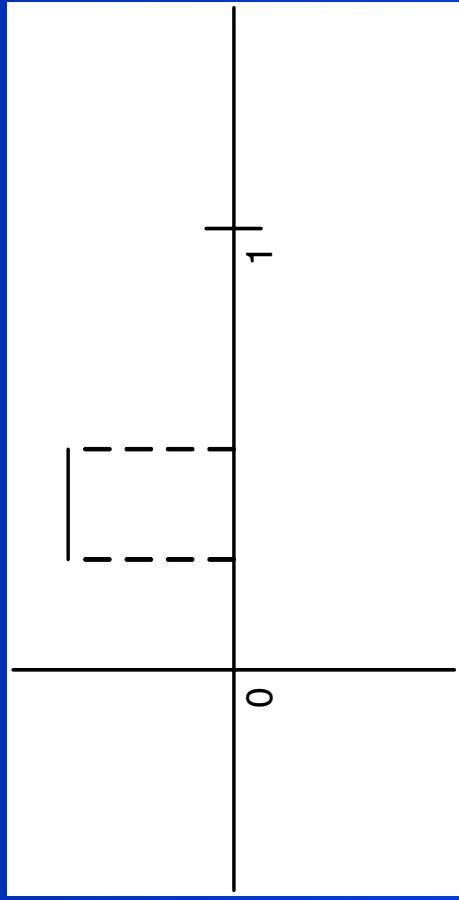
Outline of Talk

1. Motivation for multiresolution analysis: What's wrong with Fourier series?
2. Discretization of wavelets: Multirate filter banks
3. Application: The JPEG-2000 standard
4. Demonstration: The JPIP interactive client-server protocol

What's Wrong with Fourier Series for Periodic Functions?

- Fourier analysis only works on noncompact groups because of divine intervention.
- If we look closely, Fourier analysis also has problems with periodic functions:
 - ✿ No localization of information in the transform domain.
 - ✿ Problems reproducing localized singularities.

Consider the Fourier series expansion of a (periodic) step function:



$$f(x) \sim \sum c_n e^{i 2\pi n x}, \quad c_n \sim 1/n$$

What's Wrong with Fourier Series? (cont.)

- How well does this Fourier series converge?
 - ✿ Converges in $L^2([0,1])$; i.e., in the RMS sense
 - ✿ Doesn't converge uniformly on $[0,1]$
 - ✿ Converges pointwise in the intervals of continuity, uniformly on closed subintervals bounded away from the jump discontinuities
 - ✿ Converges to the average of the left and right one-sided limits at the jumps
- This is remarkable: how can a superposition of nonlocal, globally oscillatory basis functions reproduce the jump discontinuities while simultaneously converging uniformly to a constant value in intervals bounded away from the discontinuities?
- Answer: Through very, very careful phase cancellations.

What's Wrong with Fourier Series? (cont.)

- *Why on earth should we expect to get this lucky?*
 1. Complex exponentials are mutually orthogonal in $L^2([0, 1])$.
 2. An integrable function must be identically zero if all its Fourier coefficients are zero (i.e., complex exponentials are complete).
 3. Abstract nonsense (Hilbert space theory) implies that this fortuitous cancellation will occur when we sum the Fourier series in the L^2 sense.
 4. Moreover, it also follows that the Fourier transform is an L^2 -isometry:

$$\|f\|_2^2 = \sum |c_n|^2$$

- In particular, our luck hinges on L^2 theory, which begs the question:
What happens in other norms?

Fourier Series for Functions in L^p , $p < 2$

- Pick f in $L^p([0, 1])$ for $p < 2$ such that f is *not* in L^2 .
- The Hausdorff-Young inequality says that

$$\{c_n\} \in \ell^q, \|c_n\|_q \leq \|f\|_p ; \text{ where } 1/q = 1 - 1/p$$

- This inequality says the Fourier transform is continuous on L^p , but what about its inverse?
- It's possible to recover f (both pointwise and in L^p norm) by mollifying its Fourier series; e.g., using Cesaro or Abel means and taking a limit. This doesn't tell us anything about the range or the invertibility of the L^p -Fourier transform, however.
- Is there any way to get an upper bound on the p -norm of f in terms of the Fourier coefficient magnitudes, analogous to (but perhaps weaker than) the Riesz-Fischer theorem:

$$\|f\|_2^2 = \sum |c_n|^2$$

Fourier Series for Functions in L^p , $p < 2$ (cont.)

- Theorem (Paley-Zygmund, 1930): Suppose that c_n is *not* a square-summable sequence. Flip the signs of the c_n randomly to form a new (stochastic) trigonometric series:

$$\sum \pm_n c_n e^{i 2\pi n x}.$$

Then, *with probability one*, the resultant series is *not* a Fourier series of an integrable function.

- Interpretation: In sharp contrast to the L^2 case, the image under the Fourier transform of L^p , $p < 2$, is extremely sparse in ℓ^q , and there is no hope of recovering the L^p norm of a function from the magnitudes of its Fourier coefficients.
- This theorem also implies that the complex exponentials do not form an unconditional (Banach space) basis for L^p , $1 < p < 2$. It therefore appears that the phenomenon of oscillatory cancellation is much more delicate in L^p than it is in L^2 .

Littlewood-Paley Theory

- In contrast, multiresolution analysis based on the lowly Haar wavelet provides unconditional bases for all spaces $L^p([0,1])$, $1 < p < \infty$. Why are Haar wavelet expansions so robust with respect to varying the norm?
- The P-Z theorem shows that we cannot recover the L^p norm of f from its Fourier coefficient magnitudes. We can, however, recover the norm from the Fourier series using something more direct than Cesaro or Abel means. Define dyadic blocks of Fourier series components,

$$\Delta_k(x) = \sum_{n \in S_k} c_n e^{i 2\pi n x}$$

$$S_0 = \{0\}, S_k = \{n : 2^{k-1} \leq |n| < 2^k\}$$

and define the auxiliary function, γ_f , by the nonnegative series:

$$\gamma_f(x) = \left(\sum |\Delta_k(x)|^2 \right)^{1/2}$$

Littlewood-Paley Theory (cont.)

- **Theorem (Littlewood-Paley, 1931-37):** The functions f and γ_f have comparable L^p norms:

$$a_p \left\| \gamma_f \right\|_p \leq \left\| f \right\|_p \leq A_p \left\| \gamma_f \right\|_p$$

- **Interpretation:** Without actually reconstructing f , the L-P theorem tells us “how much” oscillatory cancellation is needed to recover the size (i.e., the L^p norm) of f from its Fourier series. Basically, we have to allow cancellation within dyadic frequency blocks *before* taking absolute values. This almost seems to imply that the Fourier transform is “oversampled” in the frequency domain as far as L^p approximation is concerned! Note that wavelet multiresolution approximations avoid this problem.
- Wavelet or wavelet-like exponentially scaled frequency decompositions inspired by the L-P theorem have been constructed to form “atomic decompositions” of most function spaces of interest.

Multirate Filter Banks

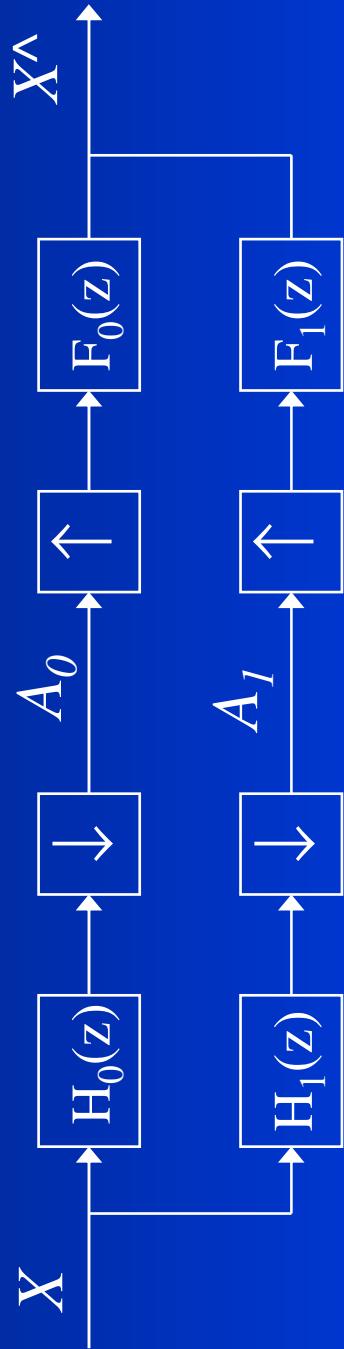
- Z-transform representation:

$$X(z) = \sum_n x(n)z^{-n}$$

- Z-transform of **vector-valued signals**:

$$\mathbf{X}(z) = \sum_n \mathbf{x}(n)z^{-n} = \begin{pmatrix} X_0(z) \\ X_1(z) \end{pmatrix}$$

2-Channel Filter Banks



Polyphase Representation

source vector: $\mathbf{X}(z) = \begin{pmatrix} X_0(z) \\ X_1(z) \end{pmatrix}$

even samples odd samples

subband vector: $\mathbf{A}(z) = \begin{pmatrix} A_0(z) \\ A_1(z) \end{pmatrix}$

lowpass highpass

analysis matrix: $\mathbf{E}(z) = \begin{bmatrix} E_{00}(z) & E_{01}(z) \\ E_{10}(z) & E_{11}(z) \end{bmatrix}$

analysis: $\mathbf{A}(z) = \mathbf{E}(z)\mathbf{X}(z)$

Synthesis: $\hat{\mathbf{X}}(z) = \mathbf{F}(z)\mathbf{A}(z)$

Perfect Reconstruction Condition

A matrix polynomial has a polynomial inverse iff:

$$|\mathbf{E}(z)| = cz^{-(N-1)} \quad (\text{PR})$$

where $c \neq 0$, $N = \text{Order}(\mathbf{E}(z))$

Lemma: Matrix polynomials satisfying condition (PR) form a group.

Cascade Lattice Factorization

Synthesize $E(z)$ as a product of elementary “building blocks”:

$$E(z) = B_0(z)B_1(z)\cdots B_m(z)$$

If the $B_i(z)$ satisfy (PR) then so does $E(z)$. Moreover, if $E(z)$ and $B(z)$ satisfy (PR) then so does $E(z)B^{-1}(z)$; i.e., factorization preserves condition (PR).

Symmetry for Odd-Length Filters

If $H_0(z)$ and $H_1(z)$ are linear phase filters of order $2N-2$ and $2N$ (resp.) then

$$\mathbf{E}(z^{-1}) = \begin{bmatrix} z^{N-1} E_{00}(z) & z^{N-2} E_{01}(z) \\ z^N E_{10}(z) & z^{N-1} E_{11}(z) \end{bmatrix} \quad (\text{SYM})$$
$$\mathbf{E}(z)\Lambda(z) = z^{-(N-1)}\Lambda(z)\mathbf{E}(z^{-1}) , \quad \Lambda(z) = \text{diag}(1, z^{-1})$$

Lemma: Matrix polynomials satisfying (SYM) form a group so synthesis & factoring preserve condition (SYM).

Low-Order Building Blocks

General solutions to equations (PR) & (SYM):

$$N = 0: \quad D_{\alpha_0, \alpha_1}(z) = \text{diag}(\alpha_0, \alpha_1) \quad ; \quad \alpha_0, \alpha_1 \neq 0$$

$$N = 1: \quad C(z) = \begin{bmatrix} 1 & 0 \\ 1 + z^{-1} & 1 \end{bmatrix} \quad (\text{mod. diag. matrices})$$

Problem: $\deg|C(z)|=0$ so can't use $C(z)$ to synthesize higher order matrix polynomials! (cf. case of even-length linear phase PR filter banks)

General Solution, $N=2$

Again, modulo diagonal constant matrices:

$$\mathbf{B}_\beta(z) = \begin{vmatrix} 1 & 1 \\ 1 + \beta z^{-1} & 1 + \beta z^{-1} + z^{-2} \end{vmatrix}, \quad \beta \neq 2$$

Since

$$|\mathbf{B}_\beta(z)| = (2 - \beta)z^{-1},$$

this solution can be used as a building block for high-order linear phase PR filter banks.

Proposed Lattice Factorization

Building blocks:

$$\mathbf{A}_{\alpha,\beta}(z) = \mathbf{B}_\beta(z) \mathbf{D}_{\alpha,1}, \quad \mathbf{A}_\alpha(z) \equiv \mathbf{A}_{\alpha,0}(z)$$

Let all but one β equal 0 in the factorization:

$$\mathbf{E}(z) = \mathbf{D}_{\alpha_0,\alpha_1} \mathbf{A}_{\alpha_2,\beta}(z) \mathbf{A}_{\alpha_3}(z) \cdots \mathbf{A}_{\alpha_N}(z)$$

Example: The Cohen-Daubechies-Feauveau 9-7 wavelet filter bank is factorable using 3 such blocks ($N=4$).

Properties of Group-Factorable Filter Banks

Theorem: Factorizations of this form are unique and have 1 multiplier per degree of freedom. Factorization algorithms follow easily from the group structure.

Theorem (Denseness of factorable filter banks): Group-factorable filter banks form an open, dense subgroup in the group of all odd-length linear phase FIR filter banks.

Complexity Comparison

$\text{MPU} \text{ (resp., APU)} = \# \text{ multiplies (resp., adds) per unit input sample.}$

Complexity of C-D-F 9-7 filter bank (used, e.g., in the FBI and JPEG-2000 standards):

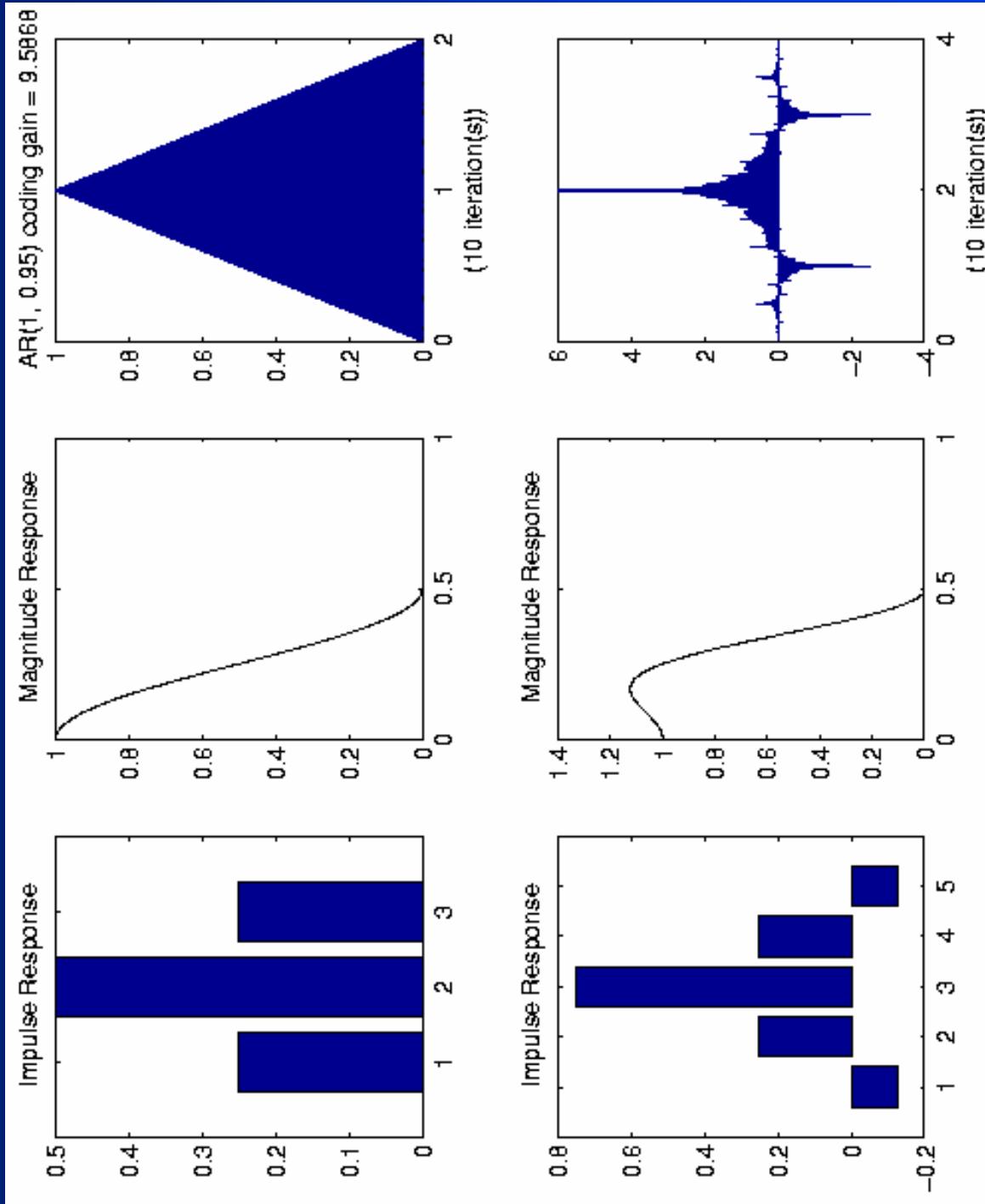
Asymptotically (for high-order filter banks), the reduction in MPU's is 50%.

	Direct Form	Lattice Structure
MPU	4.5	3
APU	7	6.5

Wavelet Filter Bank Design

- Application: Use the group structure as a parameter space to design a 7-tap/5-tap wavelet filter bank optimized for JPEG-2000 image coding.
- Goal: Produce a 7-5 filter bank that fills the 0.3 dB SNR gap between the LeGall-Tabatabai 5-3 and C-D-F 9-7 filter banks in the JPEG-2000 baseline.

LeGall-Tabatabai 5-3 Filter Bank



Design Strategy

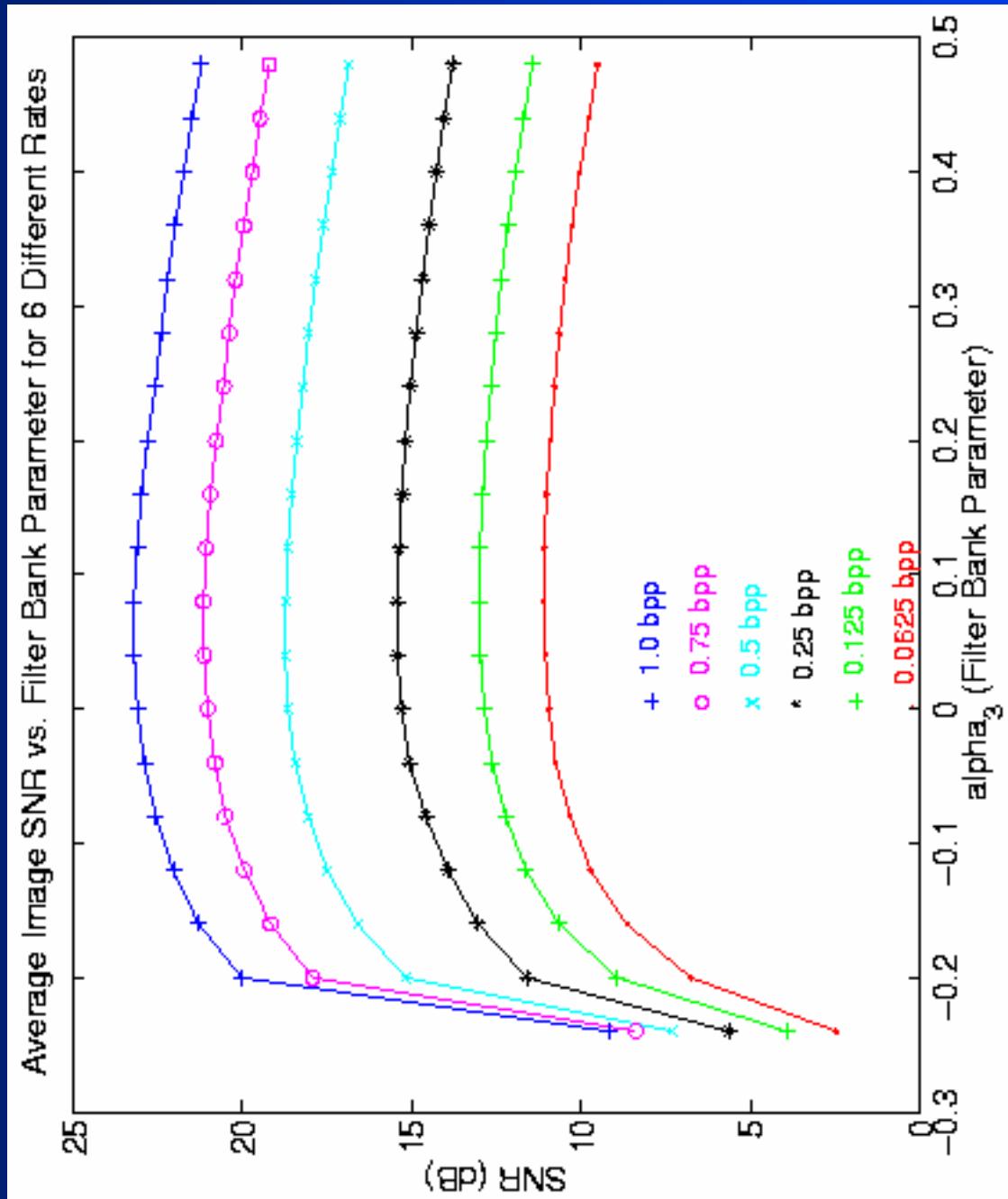
- Filter bank order is $N=3$: this means 5 parameters, 3 nontrivial degrees of freedom.

Use α_3 as a free variable and use α_2, β to impose 2 moment conditions on the filter bank.

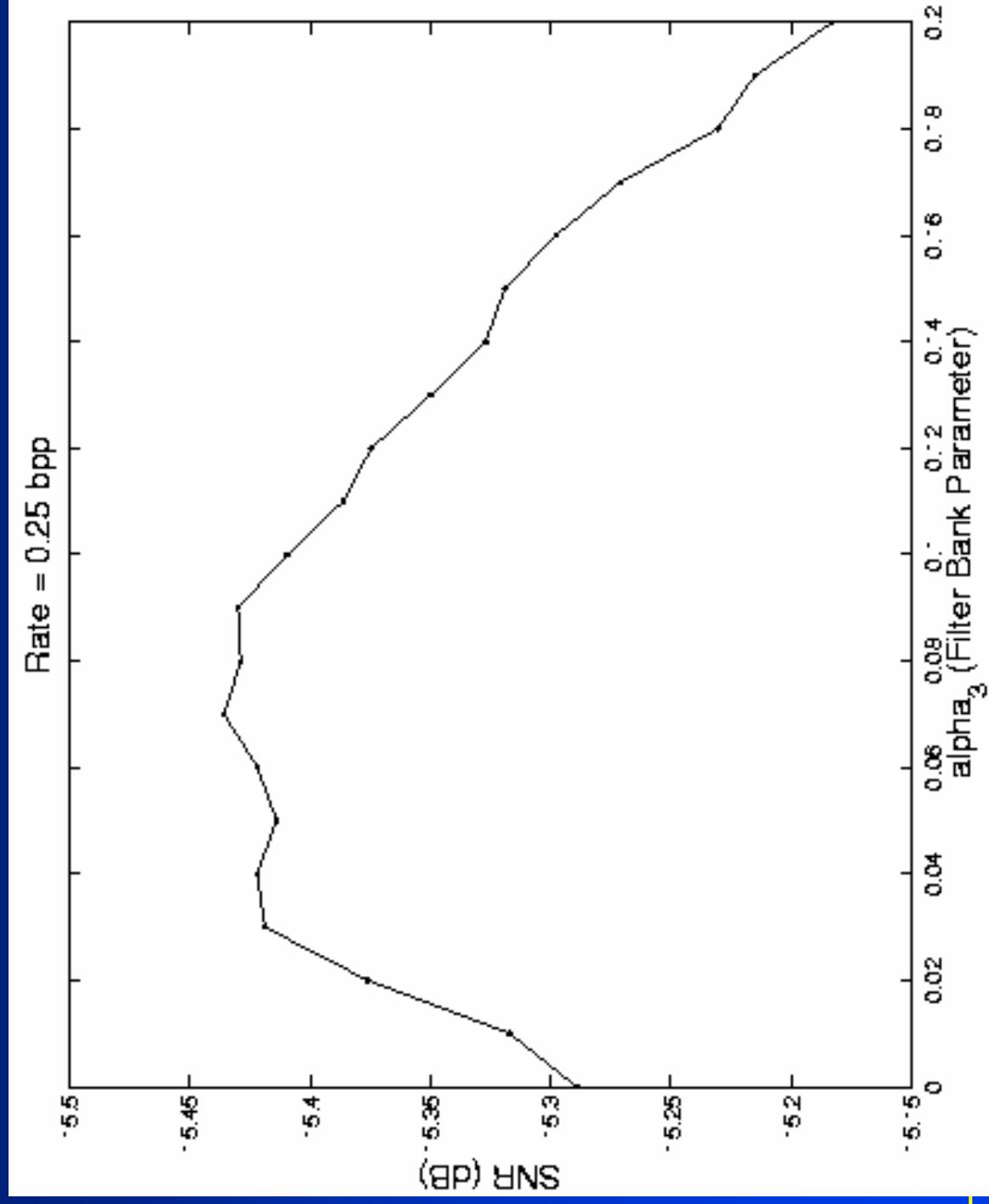
Comparison with 5-3 Filter Bank

- The “excluded” value, $Q_3 = 0$, corresponds to the natural embedding of the 5-3 filter bank in the 7-5 category as a *degenerate* 7-5 filter bank. Observe the following plots of empirical coding performance as a function of Q_3 for values close to zero. These 7-5 filter banks can be regarded as perturbations of the embedded 5-3 filter bank in this higher-dimensional filter bank manifold.

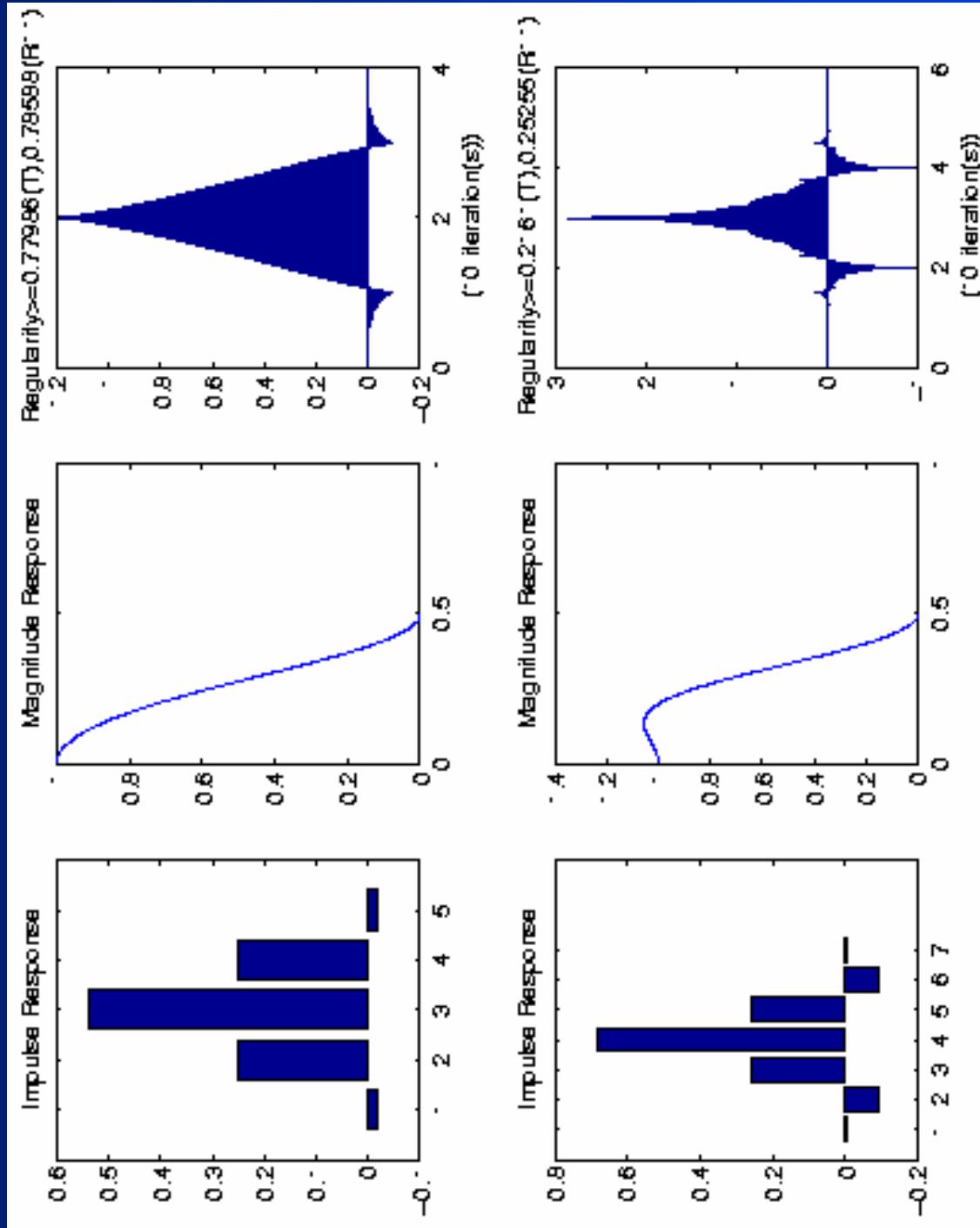
SNR Landscape: 7-5 Filter Banks



Performance at 0.25 bits/pixel



Chosen Lowpass Filters and Scaling Functions: $\alpha_3 = 0.08$



Regularity Estimates

- Using a result of O. Rioul (“Simple regularity criteria for subdivision schemes,” *SIAM J. Math. Anal.* 23 (1992), 1544-1576), we can estimate the Hölder regularity of wavelets directly from the impulse response coefficients. For the 7-5 filter bank we have designed, the exponents are estimated to be:
 - 0.786 (5-tap scaling function)
 - 0.253 (7-tap scaling function)

Regularity (cont.)

- For comparison, the LeGall-Tabatabai 5-3 wavelet family has regularity estimates:
 - 1.0 (piecewise linear B-spline)
 - 0.0 (dual scaling function)
- In a sense (“average Hoelder regularity”), the 7-5 wavelet family is slightly (“about 0.02 Hoelders”) more regular than the 5-3.

Summary: Optimal 7-5 Image Coding Filter Bank

- Improvements over 5-3 (irreversible) baseline JPEG-2000 filter bank of about 0.11 dB SNR at 1.0 bpp to about 0.15 dB SNR at 0.0625 bpp.
- This filter bank is published in JPEG-2000 Part 2, Annex G, as an optional filter bank.

n	$h_o(n)$ (lowpass analysis)	$f_o(n)$ (lowpass synthesis)	Lifting coef's: $\alpha_{0,0}$ $\alpha_{1,0}$ $\alpha_{2,0}$	K
0	79/116	27/50	$\alpha_{0,0}$ $\alpha_{1,0}$ $\alpha_{2,0}$	-175/406 609/2500
± 1	373/1450	1/4		
± 2	-21/232	-1/50		
± 3	-21/2900			25/29

What is JPEG-2000?

- It is an international standard for highly scalable coding and compression of image data, a successor to the original JPEG standard.
- As such, it defines codestreams and file formats. It is not a rendering or display tool itself, but can be used with any data visualization tool that has capabilities for reading the JPEG-2000 format.
- JPEG-2000 supports vastly expanded capabilities and scope in comparison to the original JPEG standard:
 - ✿ Very flexible progressive transmission (scalability) to enable a wide range of applications such as region-of-interest browsing.
 - ✿ Covers a much wider range of image types (much larger sizes, more pixel precisions,...) than JPEG-1, including floating point data in JP3D (Part 10).
 - ✿ Integrated support for 3-D data (multispectral, volumetric,...) including 3-D entropy coding and adaptively refined meshes (AMR data) in JP3D.
- Based on mature forms of wavelet transform and embedded arithmetic coding technology that enable hierarchical scalability and lossy-to-lossless SNR scalability, features no other coding standards possess.

What Are We Doing On This Project?

- Participating in the development of the next-generation JPEG image coding standard: JPEG-2000 (ISO 15444 and its Parts).
- This activity is proceeding in concert with widespread adoption of JPEG-2000 throughout USG agencies. NIMA is updating many of its current Mil standards with a unified JPEG-2000 profile to support imaging needs throughout the military and intelligence communities:
<http://www.ismc.nima.mil/ntb/2002SICs/00Compression%20Tutorial.htm>
- LANL team has contributed experimental results on LANL data and standards text to Part 1 (baseline JPEG-2000) and Part 2 (extensions).
- Conducted performance analysis of exploitation on hyperspectral imagery and demonstrated that common exploitation tasks are very robust with respect to compression (see slides at NIMA link).
- Actively participating in current development of Part 9 (“JP1P”): an interactive client-server protocol for JPEG-2000 imagery.
- Leading the ISO committee in current development of Part 10 (“JP3D”): extensions for 3-D data and floating point data. Brislawns has been approved by the ISO committee as Editor-In-Chief of Part 10.